



Nonrestorative sleep scale: reliable and valid for the Chinese population

S. Li¹ · D. Y. T. Fong¹ · J. Y. H. Wong¹ · K. Wilkinson² · C. Shapiro² · E. P. H. Choi¹ · B. McPherson³ · C. L. K. Lam⁴ · M. S. M. Ip⁵

Accepted: 7 February 2019 / Published online: 14 February 2019
© Springer Nature Switzerland AG 2019

Abstract

Purpose To conduct a linguistic and psychometric evaluation of a Chinese version of the Nonrestorative Sleep Scale (NRSS).

Methods The Chinese NRSS was created from a standard forward–backward translation and trialed on 10 Chinese adults. Telephone interviews were then conducted with 100 adults, who completed the Chinese NRSS, the Pittsburgh Sleep Quality Index (PSQI), the Athens Insomnia Scale (AIS), the Center for Epidemiological Studies Depression Scale (CES-D), and the Toronto Hospital Alertness Test (THAT). A household survey was conducted with 20 subjects, followed by a confirmatory factor analysis (CFA), and a bifactor model was developed to evaluate the reliability and validity of the NRSS.

Results The bifactor model had the root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), and comparative fit index (CFI) of 0.06, 0.06, and 0.97, respectively. Convergent validity was shown from the moderate associations with PSQI ($r = -0.66$, $P < 0.01$), AIS ($r = -0.65$, $P < 0.01$), CES-D ($r = -0.54$, $P < 0.01$), and THAT ($r = 0.68$, $P < 0.01$). The coefficient omega (0.92), omega hierarchical (0.81), factor determinacy (0.93), H value (0.91), explained common variance (0.63), and percentage of uncontaminated correlations (0.80) derived from the bifactor CFA supported the essential unidimensionality of NRSS.

Conclusions The Chinese NRSS is a valid and reliable essential unidimensional tool for the assessment of nonrestorative sleep in the Chinese population.

Keywords Bifactor · Confirmatory factor analysis · Nonrestorative sleep · Reliability · Validity

Introduction

Nonrestorative sleep (NRS) is generally defined as subjectively feeling unrefreshed even after sufficient sleep [1]. It is listed in the *Diagnostic and Statistical Manual of Mental Disorders* (DSM) IV [2] and the *International Classification of Sleep Disorders* 3rd edition (ICSD-3) as one of the primary defining symptoms of insomnia [3]. However, it can be a manifestation of sleep problems or disorders in itself without co-existing with other symptoms of insomnia [4]. DSM-V [5] lists NRS as “other specified insomnia disorder” or “unspecified insomnia” if its characteristics meet the standards for frequency, duration, and influence on daytime function without other sleep disorders.

NRS is more prevalent than other sleep difficulties, such as difficulty in initiating sleep [6]. It was reported that 11% of the general population in seven European countries were affected by NRS [7]. NRS may induce minor nonfatal accidents during work and leisure time [8], weaken general

✉ D. Y. T. Fong
dytfong@hku.hk

¹ School of Nursing, Li Ka Shing Faculty of Medicine, The University of Hong Kong, 21 Sassoon Road, Hong Kong, China

² Department of Psychiatry, University of Toronto, 250 College Street, Toronto, Canada

³ Division of Speech and Hearing Sciences, Faculty of Education, The University of Hong Kong, Pokfulam Road, Hong Kong, China

⁴ Department of Family Medicine and Primary Care, Li Ka Shing Faculty of Medicine, The University of Hong Kong, 21 Sassoon Road, Hong Kong, China

⁵ Department of Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, 21 Sassoon Road, Hong Kong, China

psychological well-being [9], impair quality of life [10], cause mental problems [4], and even result in suicidal ideation [11]. In addition, diabetes, chronic fatigue syndrome, fibromyalgia, and some other diseases are strongly associated with NRS [12–14]. Hence, NRS has gained increasing attention over the years as a treatment target.

Despite the clinical significance of NRS, no standardized assessment tool for NRS was available until recently [15]. NRS is associated with decreased daytime function, which involves physical, cognitive, and emotional aspects [16]. Therefore, the development of a standardized questionnaire is critical for obtaining a complete and profound understanding of NRS. In view of this, Wilkinson et al. developed the Nonrestorative Sleep Scale (NRSS) in 2013, the first instrument to assess NRS. The NRSS comprises of 12 items in four NRS domains, namely refreshment from sleep, physical/medical symptoms of NRS, daytime functioning, and affective symptoms of NRS [17]. It was developed from an original pool of 34 items after input from sleep and psychiatric experts as well as patients requiring sleep assessment [17]. The internal consistency of the scale was demonstrated, and the Cronbach's alpha for the four domains ranged from 0.64 to 0.85. Construct validity was also demonstrated with significant associations with sleepiness, insomnia, fatigue, depression, and lack of alertness. The 12-item NRSS was found to be reliable and valid for the assessment of NRS. However, the scale had not been culturally or linguistically adapted for the largest population in the world—the Chinese population.

There are known differences in sleeping habits and attitudes across various ethnic groups [18]. It has been reported that Chinese people have poorer sleep quality as compared to Western populations [19]. However, the NRS of the Chinese population has been less studied due to the lack of standardized measurements adapted for this population. This study aimed to culturally adapt the NRSS for a Chinese population that might help investigate the prevalence of NRS and enhance the epidemiological understanding of sleep issues among Chinese populations.

Methods

Participants

Eligible subjects for the study comprised of adults aged 18 years or above who could communicate in Chinese. Those not motivated to participate, unable to understand the study procedures were not recruited. Furthermore, hypertension is associated with sleep deprivation. Drugs for hypertension or sleep problems are deemed to improve sleep but how they may influence NRS is not well understood. Moreover, people under treatment of mental illness

may give inconsistent responses. Finally, the NRSS involves questions on daytime functioning that may not be relevant to people engaging in shift work. Therefore, people taking drugs for hypertension or sleep problems, being treated for mental illness, or engaged in shift work were excluded. We gathered 100 participants for a telephone survey and 20 for a household study. Most validation studies conducted previously used 2–20 subjects per item to validate their scales [20]. In view of the 12 items of the NRSS, 120 subjects were deemed reasonable.

Procedures

We conducted a telephone survey between March 14 and April 11, 2017. We obtained a publicly accessible residential telephone directory from the local government. Using simple random sampling, a set of telephone numbers was drawn out. These numbers were taken as seeds to generate another set of numbers, whose last digits were “plus or minus one or two” to capture the unlisted numbers. After removing duplicated numbers, the remaining numbers were mixed in a random order before the telephone calls were made. When there was more than one eligible person living in a household, the one with the upcoming birthday was selected. In each interview, the Cantonese-speaking research assistants read out the questions written in traditional Chinese. A total of 400 telephone numbers were dialed, and 198 phone calls obtained a response, of which 100 subjects consented to participate, resulting in a response rate of 51%. A household survey was conducted between September 2016 and July 2017 with similar selection criteria. Households were randomly selected from a representative sampling frame obtained from the Hong Kong Census and Statistics Department. When there was more than one eligible person living in the household, the one with the next birthday was asked for their consent for participation. The study protocol was approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster (Ref no.: UW16-326).

Linguistic validation of the Chinese NRSS

Following standard validation guidelines [21], two registered nurses independently translated the NRSS into traditional Chinese as used in Hong Kong and reached a consensus on the Chinese version. This version was back translated into English by an advanced practice nurse and reviewed by [KW] and [CS]. Discrepancies were identified, and minor revisions were made to clarify the meaning of the terms “physical or medical problems” (Q6) and “alert” (Q10). After further review of the Chinese NRSS by a clinician specializing in sleep [MI], the Chinese NRSS was administered

on 10 subjects to assess the clarity and relevance of the items.

Measures

Nonrestorative sleep scale (NRSS)

The original English NRSS comprises of the global scale that included all the items as well as four subscales—refreshment from sleep (3 items), physical/medical symptoms of NRS (4 items), daytime functioning (3 items), and affective symptoms (2 items). Items under the subscales of the physical/medical symptoms of NRS and the affective symptoms of NRS were negatively worded, while 10 items used the 1–10 Likert scale and the other two items used the 1–5 Likert scale [17]. A scale score was taken as the total of the corresponding item responses standardized on the 0–100 scale. For instance, the raw total of all the items ranged between 12 and 110, and the global score was $100 \times (\text{total} - 12) / (110 - 12)$.

Pittsburgh sleep quality index (PSQI)

The 19-item PSQI questionnaire assesses sleep quality during the past month [22]. This index has been used widely among general and clinical populations across countries. The items, each rated on a 0–3 scale, are grouped into 7 components. The total score ranges from 0 to 21, with a higher score indicating poorer sleep quality [22].

Athens insomnia scale (AIS)

The AIS is an 8-item self-report questionnaire that estimates sleep difficulty in the past month [23, 24]. Participants grade their sleep quality on a scale from 0 to 3. The total score ranges from 0 to 24, with a higher score corresponding to poorer sleep quality [24]. The AIS has been widely adapted in different languages with proven reliability and validity.

Center for epidemiological studies depression scale (CES-D)

The CES-D, a 20-item self-rating scale for evaluating depressive symptoms, is used to screen subjects with depression and evaluate the severity of depressive disorders [25]. Each item is graded on a 0–3 scale. The total score ranges from 0 to 60, with a higher score indicating greater severity of depression [26].

Toronto hospital alertness test (THAT)

The THAT comprises of ten self-rating items developed to estimate perceived alertness during the past week [27]. Each item uses a 6-point Likert scale in the 0–5 range, with the

last two items being inversely scored [27]. A higher total score indicates higher alertness.

Statistical analysis

The scale scores were summarized using descriptive statistics, with the floor and ceiling percentages reported to assess the scaling properties of the questionnaire. The floor and ceiling percentages of a scale indicate the percentages of participants with the lowest and highest plausible scale scores, respectively [28].

The factorial validity of the original four-scale model was assessed by a confirmatory factor analysis (CFA). The item scores were treated as continuous, since there were ten response categories in ten items and five response categories in two items. Robust maximum likelihood (MLR) estimation was used to account for any mild-to-moderate violations of normality [29]. Robust fit indices were reported. The model fit was examined by the root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), and comparative fit index (CFI) [30]. The fit was considered adequate when the RMSEA was 0.06 or below, the SRMR was 0.08 or below, and the CFI was 0.95 or higher [31].

Further psychometric evaluation of the Chinese NRSS was made using a bifactor model [32]. Coefficient omega, similar to Cronbach's α but overcoming the strong assumptions of unidimensionality and equal factor loadings of the latter, were calculated to assess internal reliability. An omega value of 0.7 or above is recommended to demonstrate a reliable total score [33]. Moreover, an omega hierarchical (omegaH) value was obtained to assess the percentage of variation attributable to a single general factor. An omegaH of at least 0.8 has been suggested to indicate reasonable unidimensionality [32]. In addition, factor score determinacy and construct replicability were evaluated by factor determinacy (FD) and H index values, respectively. An FD greater than 0.9 indicates adequate factor determinacy, and an H value greater than 0.8 indicates adequate construct replicability [32]. Furthermore, explained common variance (ECV) is recommended to be greater than 0.7 and percentage of uncontaminated correlations (PUC) greater than 0.7, the unidimensional model was evaluated as well [32]. The corrected item-scale correlations were used to assess the reliability of the Chinese NRSS. To examine convergent validity, we obtained the Spearman rank correlation coefficients of the Chinese NRSS with the PSQI, AIS, CES-D, and THAT.

Data analysis was conducted in SPSS (version 23) and RStudio-1.1.383. The package “lavaan” under RStudio was utilized to conduct the CFA [34]. The statistical indices of the bifactor model were calculated using a freely available

calculator [35]. The nominal level of significance was found to be 0.05 for all statistical tests.

Results

Demographic characteristics and NRS status of participants

A total of 120 subjects were interviewed. The characteristics are summarized in Table 1. The sample had a similar female proportion and education composition as compared to the local general population in 2016 [36]. However, the subjects were slightly older, and 46% of the subjects were retired. Table 4 summarizes the NRSS global and scale scores. There were no missing values of the NRSS. The mean global score of the NRSS was found to be 68.1 (standard deviation [SD] = 16.5). The global score had no floor and a 1% ceiling prevalence. The other scale scores had a 0–1% floor prevalence and 4–26% ceiling rate (Table 4).

Validity and reliability

Factorial validity

Table 2 displays the fit indices of various CFA models. The 4-factor model and bifactor model were satisfied fitted (Table 2). Figures 1 and 2 depict the standardized coefficients of the 4-factor and bifactor models, respectively.

Table 3 displays further statistical indices derived from the bifactor model. The coefficient omega for all scales was

Table 1 Characteristics of 120 subjects

Characteristics	<i>n</i>	%
Mean age ± SD (years, 5 missing)	56 ± 19	
Gender		
Male	54	45
Female	66	55
Educational level (1 missing)		
Primary school or below	26	22
Secondary school	51	43
Bachelor or above	42	35
Occupation (1 missing)		
Employed	27	23
Self-employed	6	5
Employer	4	3
Housewife	18	15
Seeking job	3	3
Retired	55	46
Student	6	5

SD standard deviation

Table 2 Model fit indices in confirmatory factor analysis of the Chinese NRSS (*n* = 120)

CFA model	χ^2 statistic	Degrees of freedom	RMSEA (90% CI)	SRMR	CFI
1-Factor	159.68	54	0.13 (0.11, 0.16)	0.09	0.82
4-Factor	58.09	48	0.04 (0.00, 0.08)	0.05	0.98
Bifactor	63.34	43	0.06 (0.03, 0.09)	0.06	0.97

CFA confirmatory factor analysis, RMSEA root mean square error of approximation, SRMR standardized root mean square residual, CFI comparative fit index

found to be at least 0.71. The omega hierarchical for the global scale was 0.81, and 0.10–0.44 for subscales. Only the global scale had FD and H values greater than 0.9 and 0.8. Furthermore, ECV and PUC were found to be 0.63 and 0.80, respectively.

Internal reliability

The corrected item-scale correlation of the Chinese NRSS ranged from 0.33 to 0.80 for the subscales and 0.34 to 0.76 for the global score (Table 4).

The global score was highly associated with the other scales, with correlation coefficients ranging between 0.59 and 0.84. The other scale–scale correlations were moderate, ranging from 0.32 to 0.65 (Table 5).

Convergent validity

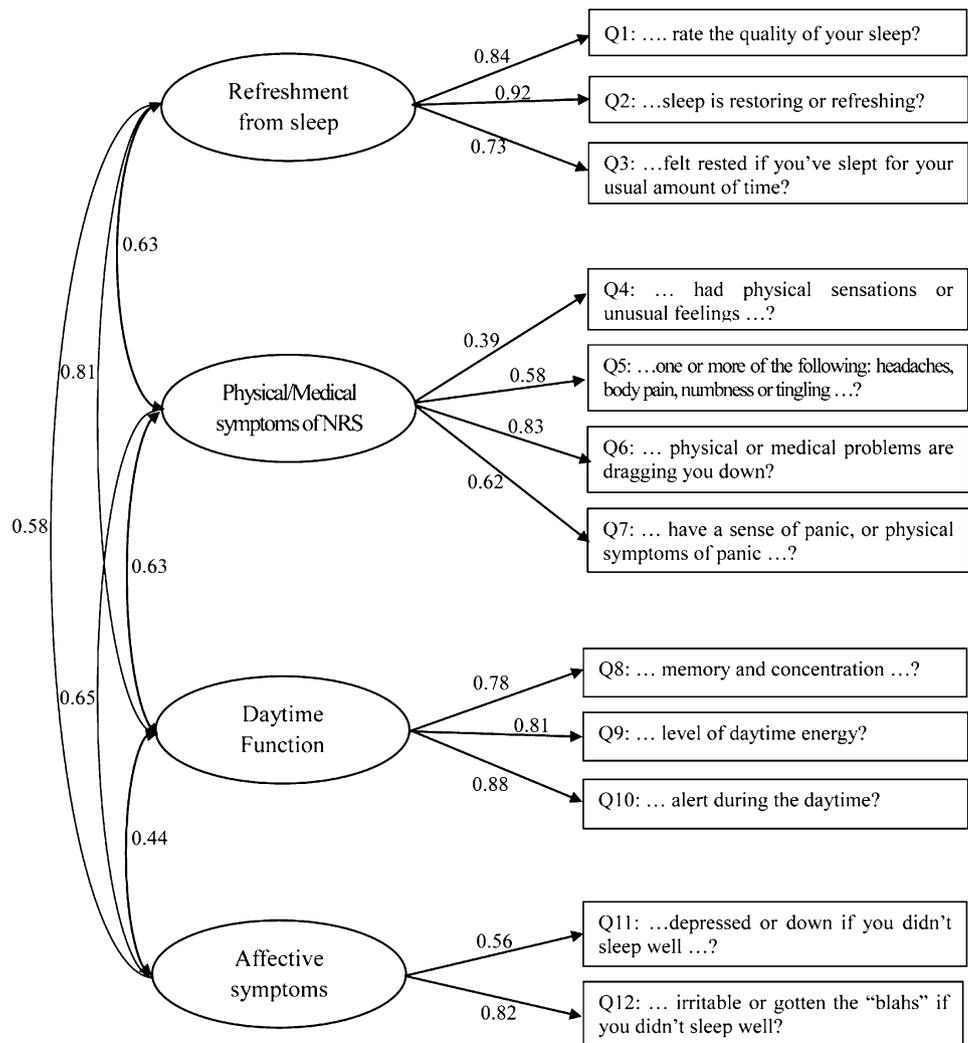
The NRSS was moderately but statistically correlated in a significant manner with the PSQI, AIS, CES-D, and THAT in the range of 0.31 to 0.68. The affective symptoms NRS scale had a relatively weak correlation of 0.31 with THAT (Table 6).

Discussion

This is the first study to assess the psychometric performance of the NRSS in a Chinese population. We rigorously translated the NRSS into Chinese and demonstrated satisfactory internal reliability and satisfactory convergent validity of the Chinese version. The bifactor model displayed the essential unidimensional nature of the NRSS for the assessment of NRS in the Chinese population, which has not previously been examined. Moreover, given the 4-factor model was appropriate, the four scales may be utilized as in the original English version.

The bifactor analysis demonstrated the essential unidimensionality of the instrument. First, the global score scale had a high omegaH of 0.81, which is merely 11% lower than

Fig. 1 Standardized coefficients of a 4-factor structure for the 12-item Chinese NRSS ($n = 120$)



its omega, whereas the group scales had a generally low omegaH. Second, only the global score had adequate FD and H values, indicating that the global score has adequate factor determinacy and construct replicability. Moreover, the ECV demonstrated that the general global score explained 63% of the common variance of the data, with the remainder attributable to the four group scales. Furthermore, the large PUC indicated that the parameter estimates of a single factor were likely to be unbiased [32]. Hence, the global score is an adequate indicator of NRS.

The corrected item-scale correlation assesses the extent to which an item is associated with its corresponding scale, and a value greater than 0.3 is necessary [37]. In this study, the corrected item-scale correlation ranged between 0.33 and 0.80. Moreover, all the Chinese NRSS scales, including the affective symptoms of NRS scale that comprises of 2 items, had satisfactory internal reliability with a coefficient omega greater than 0.7. Given the reasonable fit of the 4-factor

model, we also consider the four scales appropriate for the assessment of NRS.

The Chinese NRSS showed convergent validity demonstrating hypothesized associations with the THAT, PSQI, AIS, and CES-D. The NRSS was positively associated with the THAT but negatively correlated with the PSQI, AIS, and CES-D, which were similar to their English version [17]. There was a relatively weaker association between the Chinese NRSS and CES-D than that between the English versions. This may be due to the low depression level of our community-based sample (mean = 6.7). In contrast, the English version obtained participants from a sleep clinic, many of whom reported comorbid sleep or psychiatric conditions [38]. Nevertheless, to the best of our knowledge, the measurement invariance (MI) between the Chinese versions and English versions of the THAT, PSQI, AIS, and CES-D have not been established well. Hence, our results are only applicable to the Chinese. Whether or not similar results

Fig. 2 Standardized coefficients of a bifactor model for the 12-item Chinese NRSS ($n = 120$)

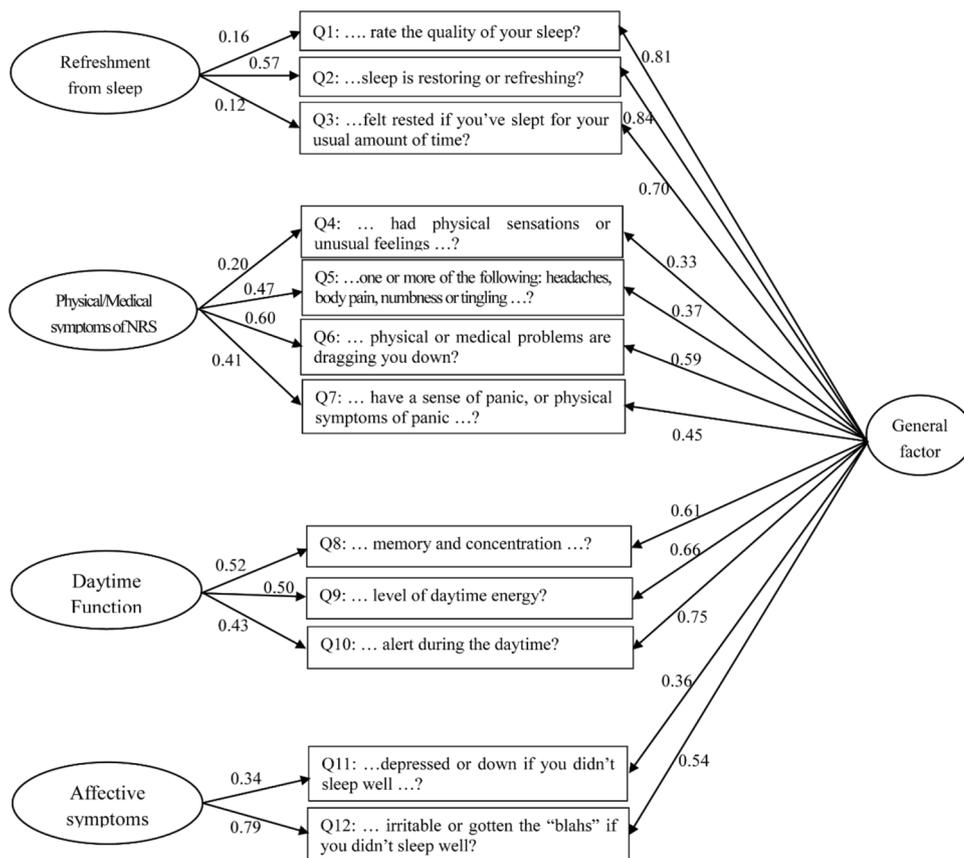


Table 3 Bifactor model statistical indices for the Chinese NRSS ($n = 120$)

Scales	Omega	Omega <i>H</i>	FD	<i>H</i>
Global score	0.92	0.81	0.93	0.91
Refreshment from sleep	0.89	0.10	0.84	0.34
Physical/medical symptoms of NRS	0.71	0.34	0.78	0.52
Daytime functioning	0.87	0.29	0.75	0.48
Affective symptoms of NRS	0.71	0.44	0.91	0.64

would be observed in other language groups requires further examination.

Despite our efforts to conduct a rigorous psychometric assessment of the Chinese NRSS, the study encountered limitations worth noting. First, our sample size of 120 may still be limited for the CFA. A study with a larger population is required to further assess the validity of the Chinese NRSS and establish the norms that would facilitate the interpretation of individual results. Second, we have not directly assessed the measurement invariance between the Chinese and English versions of NRSS, even though the conclusions regarding the factor structure of the two language versions are essentially the same. Therefore, we may validly use the

Table 4 Summary of Chinese NRSS scores ($n = 120$)

Scales (no. of items)	<i>n</i>	Mean (SD)	Min	Max	% Floor	% Ceiling	Corrected item-scale correlation
Refreshment from sleep (3)	120	63.9 (21.0)	0.0	100.0	1	6	0.67–0.80
Physical/medical symptoms of NRS (4)	120	71.0 (19.9)	12.5	100.0	0	14	0.33–0.58
Daytime functioning (3)	120	64.9 (20.0)	3.7	100.0	0	4	0.73–0.76
Affective symptoms of NRS (2)	120	71.2 (23.7)	0.0	100.0	1	26	0.46
Global score (12)	120	68.1 (16.5)	19.4	100.0	0	1	0.34–0.76

Table 5 Scale–scale correlations of the Chinese NRSS ($n = 120$)

Scales	Global score	Refreshment from sleep	Physical/medical symptoms of NRS	Daytime functioning
Refreshment from sleep	0.84**			
Physical/medical symptoms of NRS	0.82**	0.49**		
Daytime functioning	0.81**	0.65**	0.53**	
Affective symptoms of NRS	0.59**	0.46**	0.42**	0.32**

** $P < 0.01$ **Table 6** Correlation between Chinese NRSS scores and global scores of other subjective scales ($n = 100$)

Scale	PSQI	AIS	CES-D	THAT
Refreshment from sleep	−0.52**	−0.54**	−0.44**	0.55**
Physical/medical symptoms of NRS	−0.56**	−0.56**	−0.42**	0.55**
Daytime functioning	−0.47**	−0.47**	−0.38**	0.62**
Affective symptoms of NRS	−0.43**	−0.42**	−0.47**	0.31**
Global score	−0.66**	−0.65**	−0.54**	0.68**

** $P < 0.01$

NRSS in studies that administer one language version only. We did not assess the test–retest reliability and responsiveness of the NRSS in this study. Moreover, the average age in our sample was relatively high. Since the younger population is shown to have a higher risk of NRS complaints [7], the NRSS scores reported in our sample may not necessarily translate to the general population. However, given the range of NRSS scale scores that spanned across the plausible range, this would have a minimal impact on the assessment of the validity and reliability of the NRSS. Lastly, given the essential unidimensional nature of the Chinese NRSS, the possibility of reducing the length of the NRSS via the item-response theory must be explored.

In conclusion, the Chinese NRSS is a valid and reliable essential unidimensional tool for the assessment of NRS. The NRSS can help healthcare professionals, clinical experts, and the public to assess the level of NRS, which can inform the need for further treatment and health promotion.

Acknowledgements Miss Tiffany Kwok and Mr Cecil Wong, respectively, conducted the forward and backward translations, and are gratefully acknowledged.

Funding This study was financially supported by a Seed Funding for Basic Research Grant (201511159061) from The University of Hong Kong.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the insti-

tutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

- Stone, K. C., Taylor, D. J., McCrae, C. S., Kalsekar, A., & Lichstein, K. L. (2008). Nonrestorative sleep. *Sleep Medicine Reviews, 12*(4), 275–288. <https://doi.org/10.1016/j.smr.2007.12.002>.
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed). Washington, DC: American Psychiatric Association.
- Sateia, M. J. (2014). International classification of sleep disorders-third edition: Highlights and modifications. *Chest, 146*(5), 1387–1394. <https://doi.org/10.1378/chest.14-0970>.
- Sarsour, K., Van Brunt, D. L., Johnston, J. A., Foley, K. A., Morin, C. M., & Walsh, J. K. (2010). Associations of nonrestorative sleep with insomnia, depression, and daytime function. *Sleep Medicine, 11*(10), 965–972. <https://doi.org/10.1016/j.sleep.2010.08.007>.
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed). Arlington: American Psychiatric Association.
- Roth, T., Jaeger, S., Jin, R., Kalsekar, A., Stang, P. E., & Kessler, R. C. (2006). Sleep problems, comorbid mental disorders, and role functioning in the national comorbidity survey replication. *Biological Psychiatry, 60*(12), 1364–1371. <https://doi.org/10.1016/j.biopsych.2006.05.039>.
- Ohayon, M. M. (2005). Prevalence and correlates of nonrestorative sleep complaints. *Archives of Internal Medicine, 165*(1), 35–41. <https://doi.org/10.1001/archinte.165.1.35>.
- Chiu, H. Y., Wang, M. Y., Chang, C. K., Chen, C. M., Chou, K. R., Tsai, J. C., et al. (2014). Early morning awakening and nonrestorative sleep are associated with increased minor non-fatal accidents during work and leisure time. *Accident Analysis & Prevention, 71*, 10–14. <https://doi.org/10.1016/j.aap.2014.05.002>.

9. Kawada, T. (2012). Feeling refreshed by sleep can predict psychological wellbeing assessed using the general health questionnaire in male workers: A 3-year follow-up study. *Psychiatry Investigation*, 9(4), 418–421. <https://doi.org/10.4306/pi.2012.9.4.418>.
10. Liedberg, G. M., Bjork, M., & Borsbo, B. (2015). Self-reported nonrestorative sleep in fibromyalgia: Relationship to impairments of body functions, personal function factors, and quality of life. *Journal of Pain Research*, 8, 499–505. <https://doi.org/10.2147/JPR.S86611>.
11. Park, J. H., Yoo, J. H., & Kim, S. H. (2013). Associations between non-restorative sleep, short sleep duration and suicidality: Findings from a representative sample of Korean adolescents. *Psychiatry and Clinical Neurosciences*, 67(1), 28–34. <https://doi.org/10.1111/j.1440-1819.2012.02394.x>.
12. Okamoto, M., Kobayashi, Y., Nakamura, F., & Musha, T. (2017). Association between non-restorative sleep and risk of diabetes: A cross-sectional study. *Behavioural Sleep Medicine*, 15(6), 483–490. <https://doi.org/10.1080/15402002.2016.1163701>.
13. Zhang, J., Lam, S. P., Li, S. X., Li, A. M., & Wing, Y. K. (2012). The longitudinal course and impact of non-restorative sleep: A five-year community-based follow-up study. *Sleep Medicine*, 13(6), 570–576. <https://doi.org/10.1016/j.sleep.2011.12.012>.
14. Mariman, A. N., Vogelaers, D. P., Tobback, E., Delesie, L. M., Hanoulle, I. P., & Pevernagie, D. A. (2013). Sleep in the chronic fatigue syndrome. *Sleep Medicine Reviews*, 17(3), 193–199. <https://doi.org/10.1016/j.smrv.2012.06.003>.
15. Vernon, M. K., Dugar, A., Revicki, D., Treglia, M., & Buysse, D. (2010). Measurement of non-restorative sleep in insomnia: A review of the literature. *Sleep Medicine Reviews*, 14(3), 205–212. <https://doi.org/10.1016/j.smrv.2009.10.002>.
16. Wilkinson, K., & Shapiro, C. (2012). Nonrestorative sleep: symptom or unique diagnostic entity? *Sleep Medicine*, 13(6), 561–569. <https://doi.org/10.1016/j.sleep.2012.02.002>.
17. Wilkinson, K., & Shapiro, C. (2013). Development and validation of the Nonrestorative Sleep Scale (NRSS). *Journal of Clinical Sleep Medicine*, 9(9), 929–937. <https://doi.org/10.5664/jcsm.2996>.
18. National Sleep Foundation. (2010). Sleep differences among ethnic groups revealed in new poll. *ScienceDaily*. Retrieved March 27, 2010 from <https://www.sciencedaily.com/releases/2010/03/100308081740.htm>.
19. Chen, X., Wang, R., Zee, P., Lutsey, P. L., Javaheri, S., Alcantara, C., et al. (2015). Racial/ethnic differences in sleep disturbances: The multi-ethnic study of atherosclerosis (MESA). *Sleep*, 38(6), 877–888. <https://doi.org/10.5665/sleep.4732>.
20. Anthoine, E., Moret, L., Regnault, A., Sebille, V., & Hardouin, J. B. (2014). Sample size used to validate a scale: A review of publications on newly-developed patient reported outcomes measures. *Health and Quality Life Outcomes*, 12, 176. <https://doi.org/10.1186/s12955-014-0176-2>.
21. Tsang, S., Royse, C. F., & Terkawi, A. S. (2017). Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi Journal of Anaesthesia*, 11(Suppl 1), S80–S89. https://doi.org/10.4103/sja.SJA_203_17.
22. Chong, A. M. L., & Cheung, C.-K. (2012). Factor structure of a Cantonese-version pittsburgh sleep quality index. *Sleep and Biological Rhythms*, 10(2), 118–125. <https://doi.org/10.1111/j.1479-8425.2011.00532.x>.
23. Chung, K. F., Kan, K. K., & Yeung, W. F. (2011). Assessing insomnia in adolescents: Comparison of Insomnia Severity Index, Athens Insomnia Scale and Sleep Quality Index. *Sleep Medicine*, 12(5), 463–470. <https://doi.org/10.1016/j.sleep.2010.09.019>.
24. Soldatos, C. R., Dikeos, D. G., & Paparrigopoulos, T. J. (2000). Athens Insomnia Scale: Validation of an instrument based on ICD-10 criteria. *Journal of Psychosomatic Research*, 48(6), 555–560. [https://doi.org/10.1016/S0022-3999\(00\)00095-7](https://doi.org/10.1016/S0022-3999(00)00095-7).
25. Chin, W. Y., Choi, E. P., Chan, K. T., & Wong, C. K. (2015). The psychometric properties of the center for epidemiologic studies depression scale in Chinese primary care patients: factor structure, construct validity, reliability, sensitivity and responsiveness. *PLoS ONE*, 10(8), e0135131. <https://doi.org/10.1371/journal.pone.0135131>.
26. Vilagut, G., Forero, C. G., Barbaglia, G., & Alonso, J. (2016). Screening for depression in the general population with the center for epidemiologic studies depression (CES-D): A systematic review with meta-analysis. *PLoS ONE*, 11(5), e0155431. <https://doi.org/10.1371/journal.pone.0155431>.
27. Shapiro, C. M., Auch, C., Reimer, M., Kayumov, L., Heslegrave, R., Huterer, N., et al. (2006). A new approach to the construct of alertness. *Journal of Psychosomatic Research*, 60(6), 595–603. <https://doi.org/10.1016/j.jpsychores.2006.04.012>.
28. Terwee, C. B., Bot, S. D., de Boer, M. R., et al. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, 60(1), 34–42. <https://doi.org/10.1016/j.jclinepi.2006.03.012>.
29. Anastasios, S., Theodoros, A. K., Vasiliki, Y., & Konstantina, P. (2018). Using bifactor EFA, bifactor CFA and exploratory structural equation modeling to validate factor structure of the meaning in life questionnaire, Greek version. *Psychology*, 9(3), 348–371. <https://doi.org/10.4236/psych.2018.93022>.
30. Fong, D. Y., Lam, C. L., Mak, K. K., Lo, W. S., Lai, Y. K., Ho, S. Y., et al. (2010). The Short Form-12 Health Survey was a valid instrument in Chinese adolescents. *Journal of Clinical Epidemiology*, 63(9), 1020–1029. <https://doi.org/10.1016/j.jclinepi.2009.11.011>.
31. Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>.
32. Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment*, 98(3), 223–237. <https://doi.org/10.1080/00223891.2015.1089249>.
33. Gu, H. L., Wen, Z. L., & Fan, X. T. (2017). Structural validity of the Machiavellian Personality Scale: A bifactor exploratory structural equation modeling approach. *Personality and Individual Differences*, 105, 116–123. <https://doi.org/10.1016/j.paid.2016.09.042>.
34. Rosseel, Y., Oberski, D., Vanbrabant, L., Vanbrabant, L., Savalei, V., Merkle, E., et al. (2012). Package 'lavaan'. R package version 0.6-3. Retrieved from <https://cran.r-project.org/web/packages/lavaan/lavaan.n.pdf>.
35. Dueber, D. M. (2017). Bifactor Indices Calculator: A Microsoft Excel-based tool to calculate various indices relevant to bifactor CFA models. <https://doi.org/10.13023/edp.tool.01>. Retrieved from <http://sites.education.uky.edu/apslab/resources/>.
36. 2016 Population By-census Office Census and Statistics Department. (2017). 2016 Population By-census-Main Results. Hong Kong, HKSAR: Website of the Census and Statistics Department, 2017. Retrieved from <https://www.byensus2016.gov.hk/data/16bc-main-results.pdf>.
37. Nunnally, J. N. (1978). *Psychometric Theory* (2nd edn.). New York: McGraw-Hill.
38. Wilkinson, K. (2012). *The Psychometric Properties of the Non-restorative Sleep Scale and a Prospective Observational Study of the Physiological Correlates of Nonrestorative Sleep*. University of Toronto. Retrieved from https://tspace.library.utoronto.ca/bitstream/1807/32639/6/Wilkinson_Caitlin_20126_MSc_thesis.pdf.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.